

SEQUENCE LISTING

*T0640X*  
(1) GENERAL INFORMATION:

(i) APPLICANT: RADIN, DAVID N.  
CRAMER, CAROLE L.  
OISHI, KAREN K.  
WEISSENBORN, DEBORAH L.

(ii) TITLE OF INVENTION: PRODUCTION OF LYSOSOMAL ENZYMES IN  
PLANT-BASED EXPRESSION SYSTEMS

(iii) NUMBER OF SEQUENCES: 15

(iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Pennie & Edmonds  
(B) STREET: 1155 Avenue of the Americas  
(C) CITY: New York  
(D) STATE: New York  
(E) COUNTRY: USA  
(F) ZIP: 10036-2711

(v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: 08/713,928  
(B) FILING DATE: 13-SEP-1996  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 60/003,737  
(B) FILING DATE: 14-SEP-1995

(viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Coruzzi, Laura A.  
(B) REGISTRATION NUMBER: 30,742  
(C) REFERENCE/DOCKET NUMBER: 7956-0011-999

(ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: (212) 790-9090  
(B) TELEFAX: (212) 869-9741  
(C) TELEX: 66141 PENNIE

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "PCR primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTGTCTAGAG TAAGCATCAT GGCTGGC

27

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "PCR primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CACGAATTCT GGCGACGCCA CAGGTAGGTG TGA

33

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1642 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGGAGTTT CAAGTCCTTC CAGAGAGGAA TGTCCCAAGC CTTTGACTAG GGTAAGCATC	60
ATGGCTGGCA GCCTCACAGG TTTGCTTCTA CTTCAGGCAG TGTCGTGGC ATCAGGTGCC	120
CGCCCTGCA TCCCTAAAAG CTTCGGCTAC AGCTCGGTGG TGTGTGTCTG CAATGCCACA	180
TACTGTGACT CCTTGACCC CCCGACCTTT CCTGCCCTTG GTACCTTCAG CCGCTATGAG	240
AGTACACGCA GTGGCGACG GATGGGGCTG AGTATGGGGC CCATCCAGGC TAATCACACG	300
GGCACAGGCC TGCTACTGAC CCTGCAGCCA GAACAGAAAGT TCCAGAAAGT GAAGGGATT	360
GGAGGGGCCA TGACAGATGC TGCTGCTCTC AACATCCTTG CCCTGTCAACC CCCTGCCCAA	420
AATTGCTAC TTAAATCGTA CTTCTCTGAA GAAGGAATCG GATATAACAT CATCCGGTA	480
CCCATGGCCA GCTGTGACTT CTCCATCCGC ACCTACACCT ATGCAGACAC CCCTGATGAT	540
TTCCAGTTGC ACAACTTCAG CCTCCCAGAG GAAGATAACCA AGCTCAAGAT ACCCCTGATT	600
CACCGAGCCC TGCAGTTGGC CCAGCGTCCC GTTCACTCC TTGCCAGCCC CTGGACATCA	660
CCCACTGGC TCAAGACCAA TGGAGCGGTG AATGGGAAGG GGTCACTCAA GGGACAGCCC	720
GGAGACATCT ACCACCAAGAC CTGGGCCAGA TACTTGTGA AGTTCCCTGGA TGCCCTATGCT	780
GAGCACAAGT TACAGTTCTG GGCAGTGACA GCTGAAAATG AGCCTTCTGC TGGGCTGTTG	840
AGTGGATACC CCTTCCAGTG CCTGGGCTTC ACCCCTGAAC ATCAGCGAGA CTTCATTGCC	900
CGTGACCTAG GTCCTACCCCT CGCCAACAGT ACTCACCACA ATGTCCGCCT ACTCATGCTG	960
GATGACCAAC GCTTGCTGCT GCCCCACTGG GCAAAGGTGG TACTGACAGA CCCAGAAGCA	1020

GCTAAATATG TTCATGGCAT TGCTGTACAT TGGTACCTGG ACTTTCTGGC TCCAGCCAA	1080
GCCACCCCTAG GGGAGACACA CCGCCTGTTG CCCAACACCA TGCTCTTGCG CTAGAGGCC	1140
TGTGTGGGCT CCAAGTTCTG GGAGCAGAGT GTGCGGCTAG GCTCCTGGGA TCGAGGGATG	1200
CAGTACAGCC ACAGCATCAT CACGAACCTC CTGTACCATG TGGTCGGCTG GACCGACTGG	1260
AACCTTGCCC TGAACCCCGA AGGAGGACCC AATTGGGTGC GTAACTTGT CGACAGTCCC	1320
ATCATTGTAG ACGTCACCAG GGACACGTTT TACAAACAGC CCATGTTCTA CCACCTTGGC	1380
CACTTCAGCA AGTTCATTCC TGAGGGCTCC CAGAGAGTGG GGCTGGTGC CAGTCAGAAG	1440
AACGACCTGG ACGCAGTGGC ACTGATGCAT CCCGATGGCT CTGCTGTTGT GGTCGTGCTA	1500
AACCGCTCCT CTAAGGATGT GCCTCTTACC ATCAAGGATC CTGCTGTGGG CTTCCTGGAG	1560
ACAATCTCAC CTGGCTACTC CATTACACACC TACCTGTGGC GTCGCCAGAA TTCGGACTAC	1620
AAGGACGACG ATGACAAGTT GA	1642

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 546 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Glu Phe Ser Ser Pro Ser Arg Glu Glu Cys Pro Lys Pro Leu Ser	
1 5 10 15	
Arg Val Ser Ile Met Ala Gly Ser Leu Thr Gly Leu Leu Leu Leu Gln	
20 25 30	
Ala Val Ser Trp Ala Ser Gly Ala Arg Pro Cys Ile Pro Lys Ser Phe	
35 40 45	
Gly Tyr Ser Ser Val Val Cys Val Cys Asn Ala Thr Tyr Cys Asp Ser	
50 55 60	
Phe Asp Pro Pro Thr Phe Pro Ala Leu Gly Thr Phe Ser Arg Tyr Glu	
65 70 75 80	
Ser Thr Arg Ser Gly Arg Arg Met Glu Leu Ser Met Gly Pro Ile Gln	
85 90 95	
Ala Asn His Thr Gly Thr Gly Leu Leu Leu Thr Leu Gln Pro Glu Gln	
100 105 110	
Lys Phe Gln Lys Val Lys Gly Phe Gly Gly Ala Met Thr Asp Ala Ala	
115 120 125	
Ala Leu Asn Ile Leu Ala Leu Ser Pro Pro Ala Gln Asn Leu Leu Leu	
130 135 140	
Lys Ser Tyr Phe Ser Glu Glu Gly Ile Gly Tyr Asn Ile Ile Arg Val	
145 150 155 160	

Pro Met Ala Ser Cys Asp Phe Ser Ile Arg Thr Tyr Thr Tyr Ala Asp  
165 170 175

Thr Pro Asp Asp Phe Gln Leu His Asn Phe Ser Leu Pro Glu Glu Asp  
180 185 190

Thr Lys Leu Lys Ile Pro Leu Ile His Arg Ala Leu Gln Leu Ala Gln  
195 200 205

Arg Pro Val Ser Leu Leu Ala Ser Pro Trp Thr Ser Pro Thr Trp Leu  
210 215 220

Lys Thr Asn Gly Ala Val Asn Gly Lys Gly Ser Leu Lys Gly Gln Pro  
225 230 235 240

Gly Asp Ile Tyr His Gln Thr Trp Ala Arg Tyr Phe Val Lys Phe Leu  
245 250 255

Asp Ala Tyr Ala Glu His Lys Leu Gln Phe Trp Ala Val Thr Ala Glu  
260 265 270

Asn Glu Pro Ser Ala Gly Leu Leu Ser Gly Tyr Pro Phe Gln Cys Leu  
275 280 285

Gly Phe Thr Pro Glu His Gln Arg Asp Phe Ile Ala Arg Asp Leu Gly  
290 295 300

Pro Thr Leu Ala Asn Ser Thr His His Asn Val Arg Leu Leu Met Leu  
305 310 315 320

Asp Asp Gln Arg Leu Leu Leu Pro His Trp Ala Lys Val Val Leu Thr  
325 330 335

Asp Pro Glu Ala Ala Lys Tyr Val His Gly Ile Ala Val His Trp Tyr  
340 345 350

Leu Asp Phe Leu Ala Pro Ala Lys Ala Thr Leu Gly Glu Thr His Arg  
355 360 365

Leu Phe Pro Asn Thr Met Leu Phe Ala Ser Glu Ala Cys Val Gly Ser  
370 375 380

Lys Phe Trp Glu Gln Ser Val Arg Leu Gly Ser Trp Asp Arg Gly Met  
385 390 395 400

Gln Tyr Ser His Ser Ile Ile Thr Asn Leu Leu Tyr His Val Val Gly  
405 410 415

Trp Thr Asp Trp Asn Leu Ala Leu Asn Pro Glu Gly Gly Pro Asn Trp  
420 425 430

Val Arg Asn Phe Val Asp Ser Pro Ile Ile Val Asp Val Thr Lys Asp  
435 440 445

Thr Phe Tyr Lys Gln Pro Met Phe Tyr His Leu Gly His Phe Ser Lys  
450 455 460

Phe Ile Pro Glu Gly Ser Gln Arg Val Gly Leu Val Ala Ser Gln Lys  
465 470 475 480

Asn Asp Leu Asp Ala Val Ala Leu Met His Pro Asp Gly Ser Ala Val  
485 490 495

Val Val Val Leu Asn Arg Ser Ser Lys Asp Val Pro Leu Thr Ile Lys  
500 505 510

Asp Pro Ala Val Gly Phe Leu Glu Thr Ile Ser Pro Gly Tyr Ser Ile

515

520

525

His Thr Tyr Leu Trp Arg Arg Gln Asn Ser Asp Tyr Lys Asp Asp Asp  
 530 535 540

Asp Lys  
 545

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 471 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "MeGA Promoter"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CAATACGATA TTACCGAATA TTATACTAAA TCAAAATTAA ATTATCATA TCGAATTATT	60
AAACTGATAT TTCAAATTAA AATATTAAAT ATCTACTTTC AACTATTATT ACCTAATTAT	120
CAAATGCAAA ATGTATGAGT TATTCATAA TAGCCGAGT TCGTATCCAA ATATTTACA	180
CTTGACCAGT CAACTTGACT ATATAAAACT TTACTTCAAA AAATTAAAAAA AAAAAGAAAAG	240
TATATTATTG TAAAAGATAA TACTCCATTC AAAATATAAA ATGAAAAAAAG TCCAGCGCGG	300
CAACCGGGTT CCTCTATAAA TACATTTCTC ACATCTTCTC TTCTCCTCAC ATCCCATCAC	360
TCTTCTTTA ACAATTATAC TTGTCAATCA TCAATCCCAC AAACAACACT TTTTCTCTCC	420
TCTTTTCCT CACCGCGGC AGACTTACCG GTGAAATCTA GAGTAAGCAT C	471

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR primer"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CTAGTCTAGA ATGCGTCCCC TGCGCCCCCG CG	32
-------------------------------------	----

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "PCR primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGAATTCGAG CTCTCATGGA TTGCCCGGGG ATG

33

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2067 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATGCGTCCCC	TGCGCCCCCG	CGCCGCGCTG	CTGGCGCTCC	TGGCCTCGCT	CCTGGCCGCG	60
CCCCCGGTGG	CCCCGGCCGA	GGCCCCGCAC	CTGGTGCAGG	TGGACGCGGC	CCGCGCGCTG	120
TGGCCCCCTGC	GGCGCTTCTG	GAGGAGCACCA	GGCTTCTGCC	CCCCGCTGCC	ACACAGCCAG	180
GCTGACCAGT	ACGTCCCTCAG	CTGGGACCAAG	CAGCTCAACC	TCGCCTATGT	GGGCGCCGTC	240
CCTCACCGCG	GCATCAAGCA	GGTCCGGACC	CACTGGCTGC	TGGAGCTTGT	CACCACCAGG	300
GGGTCCACTG	GACGGGGCCT	GAGCTACAAC	TTCACCCACC	TGGACGGGTA	CTTGGACCTT	360
CTCAGGGAGA	ACCAGCTCCT	CCCAGGGTTT	GAGCTGATGG	GCAGCGCCCTC	GGGCCACTTC	420
ACTGACTTTG	AGGACAAGCA	GCAGGTGTTT	GAGTGGAAAGG	ACTTGGTCTC	CAGCCTGGCC	480
AGGAGATA	TCGGTAGGTA	CGGACTGGCG	CATGTTCCA	AGTGGAACTT	CGAGACGTGG	540
AATGAGCCAG	ACCACCAACGA	CTTGACAAC	GTCTCCATGA	CCATGCAAGG	CTTCCTGAAC	600
TACTACGATG	CCTGCTCGGA	GGGTCTGCGC	GCCGCCAGCC	CCGCCCTGCG	GCTGGGAGGC	660
CCCGCGACT	CCTTCCACAC	CCCACCGCGA	TCCCCGCTGA	GCTGGGGCCT	CCTGCGCCAC	720
TGCCACGACG	GTACCAA	CTTCACTGGG	GAGGCGGGCG	TGCGGCTGGA	CTACATCTCC	780
CTCCACAGGA	AGGGTGGCG	CAGCTCCATC	TCCATCCTGG	AGCAGGAGAA	GGTCGTCGCG	840
CACGAGATCC	GGCAGCTCTT	CCCCAAGTTC	GCGGACACCC	CCATTACAA	CGACGAGGCG	900
GACCCGCTGG	TGGGCTGGTC	CCTGCCACAG	CCGTGGAGGG	CGGACGTGAC	CTACGCGGCC	960
ATGGTGGTGA	AGGTCA	TGCGCTAC	TGGCCAACAC	CACCTCCGCC	1020	
TTCCCCCTACG	CGCTCCTGAG	CAACGACAAT	GCCTTCCTGA	GCTACCACCC	GCACCCCTTC	1080
GCGCAGCGCA	CGCTCACCGC	GCGCTTCCAG	GTCAACAACA	CCCGCCCGCC	GCACGTGCA	1140
CTGTTGCGCA	AGCCGGTGCT	CACGCCATG	GGGCTGCTGG	CGCTGCTGGA	TGAGGAGCAG	1200
CTCTGGGCCG	AAAGTGTGCA	GGCCGGGACC	GTCCTGGACA	GCAACCACAC	GGTGGCGTC	1260

CTGGCCAGCG	CCCACCGCCC	CCAGGGCCCG	GCCGACGCCT	GGCGGCCGC	GGTGCTGATC	1320
TACGCGAGCG	ACGACACCCG	CGCCCACCCC	AACCGCAGCG	TCGCGGTGAC	CCTGCGGCTG	1380
CGCGGGGTGC	CCCCCGGCC	GGGCCTGGTC	TACGTCACGC	GCTACCTGGA	CAACGGGCTC	1440
TGCAGCCCCG	ACGGCGAGTG	GCGGCGCCTG	GGCCGGCCCG	TCTTCCCCAC	GGCAGAGCAG	1500
TTCCGGCGCA	TGCGCGCGGC	TGAGGACCCG	GTGGCCGCGG	CGCCCCGCC	CTTACCCGCC	1560
GGCGGCCGCC	TGACCCTGCG	CCCCGCGCTG	CGGCTGCCGT	CGCTTTGCT	GGTGCACGTG	1620
TGTGCGCGCC	CCGAGAAGCC	GCCCAGGCAG	GTCACGCGC	TCCGCGCCCT	GCCCCTGACC	1680
CAAGGGCAGC	TGGTTCTGGT	CTGGTCGGAT	GAACACGTGG	GCTCCAAGTG	CCTGTGGACA	1740
TACGAGATCC	AGTTCTCTCA	GGACGGTAAG	GCGTACACCC	CGGTCAGCAG	GAAGCCATCG	1800
ACCTCAACC	TCTTGTGTT	CAGCCCAGAC	ACAGGTGCTG	TCTCTGGCTC	CTACCGAGTT	1860
CGAGCCCTGG	ACTACTGGC	CCGACCAGGC	CCCTCTCGG	ACCCTGTGCC	GTACCTGGAG	1920
GTCCCTGTGC	CAAGAGGGCC	CCCATCCCCG	GGCAATCCAT	GAGCCTGTGC	TGAGCCCCAG	1980
TGGGTTGCAC	CTCCACCGGC	AGTCAGCGAG	CTGGGGCTGC	ACTGTGCCCA	TGCTGCCCTC	2040
CCATCACCCC	CTTTGCAATA	TATTTTT				2067

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 653 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met	Arg	Pro	Leu	Arg	Pro	Arg	Ala	Ala	Leu	Leu	Ala	Leu	Leu	Ala	Ser
1					5				10					15	
Leu	Leu	Ala	Ala	Pro	Pro	Val	Ala	Pro	Ala	Glu	Ala	Pro	His	Leu	Val
					20			25				30			
His	Val	Asp	Ala	Ala	Arg	Ala	Leu	Trp	Pro	Leu	Arg	Arg	Phe	Trp	Arg
					35			40			45				
Ser	Thr	Gly	Phe	Cys	Pro	Pro	Leu	Pro	His	Ser	Gln	Ala	Asp	Gln	Tyr
					50			55			60				
Val	Leu	Ser	Trp	Asp	Gln	Gln	Leu	Asn	Leu	Ala	Tyr	Val	Gly	Ala	Val
					65			70			75		80		
Pro	His	Arg	Gly	Ile	Lys	Gln	Val	Arg	Thr	His	Trp	Leu	Leu	Glu	Leu
					85			90			95				
Val	Thr	Thr	Arg	Gly	Ser	Thr	Gly	Arg	Gly	Leu	Ser	Tyr	Asn	Phe	Thr
					100			105			110				
His	Leu	Asp	Gly	Thr	Leu	Asp	Leu	Leu	Arg	Glu	Asn	Gln	Leu	Leu	Pro
					115			120			125				

Gly Phe Glu Leu Met Gly Ser Ala Ser Gly His Phe Thr Asp Phe Glu  
130 135 140

Asp Lys Gln Gln Val Phe Glu Trp Lys Asp Leu Val Ser Ser Leu Ala  
145 150 155 160

Arg Arg Tyr Ile Gly Arg Tyr Gly Leu Ala His Val Ser Lys Trp Asn  
165 170 175

Phe Glu Thr Trp Asn Glu Pro Asp His His Asp Phe Asp Asn Val Ser  
180 185 190

Met Thr Met Gln Gly Phe Leu Asn Tyr Tyr Asp Ala Cys Ser Glu Gly  
195 200 205

Leu Arg Ala Ala Ser Pro Ala Leu Arg Leu Gly Gly Pro Gly Asp Ser  
210 215 220

Phe His Thr Pro Pro Arg Ser Pro Leu Ser Trp Gly Leu Leu Arg His  
225 230 235 240

Cys His Asp Gly Thr Asn Phe Phe Thr Gly Glu Ala Gly Val Arg Leu  
245 250 255

Asp Tyr Ile Ser Leu His Arg Lys Gly Ala Arg Ser Ser Ile Ser Ile  
260 265 270

Leu Glu Gln Glu Lys Val Val Ala Gln Glu Ile Arg Gln Leu Phe Pro  
275 280 285

Lys Phe Ala Asp Thr Pro Ile Tyr Asn Asp Glu Ala Asp Pro Leu Val  
290 295 300

Gly Trp Ser Leu Pro Gln Pro Trp Arg Ala Asp Val Thr Tyr Ala Ala  
305 310 315 320

Met Val Val Lys Val Ile Ala Gln His Gln Asn Leu Leu Leu Ala Asn  
325 330 335

Thr Thr Ser Ala Phe Pro Tyr Ala Leu Leu Ser Asn Asp Asn Ala Phe  
340 345 350

Leu Ser Tyr His Pro His Pro Phe Ala Gln Arg Thr Leu Thr Ala Arg  
355 360 365

Phe Gln Val Asn Asn Thr Arg Pro Pro His Val Gln Leu Leu Arg Lys  
370 375 380

Pro Val Leu Thr Ala Met Gly Leu Leu Ala Leu Leu Asp Glu Glu Gln  
385 390 395 400

Leu Trp Ala Glu Val Ser Gln Ala Gly Thr Val Leu Asp Ser Asn His  
405 410 415

Thr Val Gly Val Leu Ala Ser Ala His Arg Pro Gln Gly Pro Ala Asp  
420 425 430

Ala Trp Arg Ala Ala Val Leu Ile Tyr Ala Ser Asp Asp Thr Arg Ala  
435 440 445

His Pro Asn Arg Ser Val Ala Val Thr Leu Arg Leu Arg Gly Val Pro  
450 455 460

Pro Gly Pro Gly Leu Val Tyr Val Thr Arg Tyr Leu Asp Asn Gly Leu  
465 470 475 480

Cys Ser Pro Asp Gly Glu Trp Arg Arg Leu Gly Arg Pro Val Phe Pro

485	490	495
Thr Ala Glu Gln Phe Arg Arg Met Arg Ala Ala Glu Asp Pro Val Ala		
500	505	510
Ala Ala Pro Arg Pro Leu Pro Ala Gly Gly Arg Leu Thr Leu Arg Pro		
515	520	525
Ala Leu Arg Leu Pro Ser Leu Leu Val His Val Cys-Ala Arg Pro		
530	535	540
Glu Lys Pro Pro Gly Gln Val Thr Arg Leu Arg Ala Leu Pro Leu Thr		
545	550	560
Gln Gly Gln Leu Val Leu Val Trp Ser Asp Glu His Val Gly Ser Lys		
565	570	575
Cys Leu Trp Thr Tyr Glu Ile Gln Phe Ser Gln Asp Gly Lys Ala Tyr		
580	585	590
Thr Pro Val Ser Arg Lys Pro Ser Thr Phe Asn Leu Phe Val Phe Ser		
595	600	605
Pro Asp Thr Gly Ala Val Ser Gly Ser Tyr Arg Val Arg Ala Leu Asp		
610	615	620
Tyr Trp Ala Arg Pro Gly Pro Phe Ser Asp Pro Val Pro Tyr Leu Glu		
625	630	640
Val Pro Val Pro Arg Gly Pro Pro Ser Pro Gly Asn Pro		
645	650	

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Asp Tyr Lys Asp Asp Asp Asp Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "PCR primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCCTATGCTG AGCACAAAGTT ACAG

24

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "Complementary sequence of a PCR primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TCCCTTGAGC TCGTCACTGG CGACGCCACA GGTA

34

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 48 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CAGAATTCCG ACTACAAGGA CGACGATGAC AAGTAGGAGC TCGAATT

48

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

*B Coal* (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Asn Ser Asp Tyr Lys Asp Asp Asp Asp Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 4 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Lys Asp Glu Leu  
1